1-1-2009

Goal Attainment Scaling to Determine Effectiveness of School Psychology Practicum Students

Kristen Frace

Goal Attainment Scaling to Determine Effectiveness of School Psychology Practicum Students

Thesis submitted to the
Graduate College of
Marshall University

In partial fulfillment of
the requirements for the degree of
Education Specialist
In School Psychology

by

Kristen Frace

Sandra S. Stroebel, Ph.D., Committee Chairperson
Stephen O'Keefe, Ph.D.
R. Vernon Haning, M.D.

Marshall University

August 2009

ABSTRACT

Goal Attainment Scale to Determine Effectiveness of School Psychology Practicum Students

The Goal Attainment Scaling (GAS) is a system used by school psychologists and other professionals to evaluate the effectiveness of their interventions and services they provide by assessing student outcome data. The Marshall University Summer Enrichment Program (MUSEP) utilized the GAS to determine the effectiveness of the school psychology practicum students and its program. This study looks at the effectiveness of graduate students within seven teams and the services they provided to children during the 2008 MUSEP. Results of the t-test indicated that 74% of children made significant gains in their specified academic and behavior goals. A UNIANOVA and a Kruskal-Wallis test were run with results indicating there were no statistically significant differences between teams.

ACKNOWLEDGEMENTS

First, I would like to express my appreciation to my family, friends, and loved ones for their continuous support, encouragement and inspiration throughout this study as well as the School Psychology Program. They have helped me to follow through with my goal of becoming a school psychologist. I would like to express my sincere gratitude and appreciation to Dr. Sandra Stroebel for her continuous guidance, support, patience, and insight. Her encouragement and dedication of time made it possible for me to complete this study. Thank you for believing in me. I would like to express my gratitude to Dr. Stephen O'Keefe who has provided continuous guidance as well as a wealth of knowledge I will take with me and apply as I begin my career as a school psychologist. He is truly amazing and is an invaluable asset to the field of School Psychology. I would also like to thank Dr. R. Vernon Haning for dedicating his time to help me with the data analysis; as well develop a better understanding of statistical methods. His time and dedication to me during this study is greatly appreciated. I would also like to extend my appreciation to my teammate in the 2008 MUSEP, Ashley Barr, who guided me to have a better understanding of the Goal Attainment Scale and also demonstrated characteristics of a true school psychologist.

TABLE OF CONTENTS

LIST OF TABLES

Goal Attainment Scaling to Determine Effectiveness of School Psychology Practicum Students

CHAPTER I

Evaluation of practicum students, interns, and practicing School Psychologists is required by the National Association of School Psychologists (NASP, 2000). These evaluations take many forms yet one purpose is to determine the effectiveness of interventions and services that are provided to children, teens and parents. A relatively new instrument, the Goal Attainment Scale (GAS) (Kiresuk, Smith, & Cardillo, 1994) has been proposed as a measure used to gain insight into the effectiveness of services provided. School Psychologists can use GAS as a way to evaluate progress of goals they have implemented for students.

National Association of School Psychologists' Requirements

The National Association of School Psychologists (NASP) strongly promotes the professional evaluation of school psychologists and interns as a means of assuring effective practices to support the educational attainment of all children (NASP, 2004). Supervision and evaluation in school psychology focuses on promoting effective growth and exemplary professional practice leading to improved performance by all, including the school psychologist, supervisor, students, and the entire school community (NASP, 2000). NASP believes that supervisors can promote continuous professional development in order to improve practice for school psychologists. Supervisors and school psychologists should ensure that the evaluation of service units is both summative and formative, and based upon specific, measurable goals (NASP, 2000). NASP urges evaluation to ensure that supervisors provide accessible, constructive support, monitoring and feedback to practitioners in a manner that promotes professional development and effective service delivery (NASP, 2000). NASP (2004) guidelines indicate that the level and type of supervision must be adequate to ensure effective and accountable services. To ensure this, the supervision process should be continuous, positive, systematic, and collaborative. NASP has also developed guidelines for evaluation of school

psychologists which holds supervisors accountable for having a plan that includes specific, measurable goals.

Evaluation of Practicing School Psychologists

Chafouleas, Clonan & Vanakuen (2002) evaluated supervision and evaluation practices of school psychologists. The study included a national survey of a random sample of 189 nationally certified school psychologists.

Evaluation was defined as a determination of the significance of an individual's professional skills as well as assessment which includes appraisal of strengths as well as skills in need of improvement (Chafouleas et al., 2002). Questions pertaining to the area of evaluation included who was responsible for conducting the evaluation, what the evaluation entailed, the purpose for evaluation, and what the respondent would like to change about the evaluation process. Results of the evaluation practices indicated that administrators were most often responsible for evaluating the school psychologist. Sixty eight percent of respondents reported that written criteria for their evaluation was provided to them at or prior to the actual evaluation. Of those provided with criteria, only 45% were evaluated using criteria specifically designed for evaluating a school psychologist. The most common responses for purposes of evaluation included a combination of placement in the permanent file, to document work completed, and to discuss professional development. The recommendations for the evaluation process included having a school psychologist or someone familiar with the roles of a school psychologist to conduct the evaluation rather than a district-level or building-level administrator, and to use criteria specifically designed for the evaluation of school psychologists. Overall, the results of the study suggest that school psychologists may find evaluation more professionally rewarding when provided with more regular and formal contacts, and when someone knowledgeable with school psychology is involved (Chafouleas et al., 2002).

Evaluation Best Practices

Supervisors and evaluators of school psychologists need to ensure that evaluation and monitoring is provided. Knowing where school psychologists' skills are, where their skills need to be, and how far they need to go to obtain the desired performance targets is imperative in the supervision and evaluation process. Supervisors need to set goals, implement plans for change, monitor progress, and make decisions about progress (Allison, 2002). With these goals in mind, Heartland Area Education Agency 11 developed a rating scale to evaluate school psychologists (Allison, 2002). The rating scale includes: C, competent: consistently meets the requirements of the assignment; NI, needs improvement: growth needed to meet the requirements of the assignment; U, unsatisfactory: does not meet the requirements of the assignment. This rating scale can be used to determine the effectiveness of the particular assignments and job responsibilities required of the school psychologist. Evaluation by supervisors with rating scales such as this one provides a method of performance feedback and monitoring of skill sets for psychologists, shows that specific elements of practice are imperative, and allows for the supervisor to evaluate the psychologists' attainment and demonstration of skills in the field (Allison, 2002).

Gibbons and Shinn (2002) and Gibbons and Silberglitt (2008) found that tabulation, time analyses, and surveys have been the most widely used methods to evaluate the effectiveness of services that school psychologists are providing. Time analyses and tabulation are collected by having school psychologists keep a log where a description of the activities and services they provide, along with the amount of time taken for each is documented. School psychologists can also use surveys as a way to measure their effectiveness and gain feedback. Beyond time analyses and surveys, school psychologists can evaluate their effectiveness using student outcome data and look at student improvement by which services have been provided. Outcome data is intended to provide information about the changes that are being measured either academically or

behaviorally. Fuchs (2002) stresses that clearly written justifiable goals and procedures for evaluating goal attainment are potentially the key factors for obtaining student outcome data.

Goal Attainment Scaling

Goal Attainment Scaling (GAS) is an outcome measure that was originally developed by Kiresuk and Sherman (1968) for use in the mental health field. In 1969, the National Institute of Mental Health provided funding to develop, implement, and disseminate GAS (Kiresuk, Smith, & Cardillo, 1994). Evidenced-based research for the use of the GAS as it pertains to the effectiveness of outcome data and established goals has been documented (Kiresuk et al., 1994). While most of the original research was in mental health, it has also been applied to education. In education, the method is used to monitor improvement in skills and knowledge in programs (Kiresuk et al., 1994).

While conducting research, Kiresuk et al. (1994) found that a distinguishing feature of the GAS is its adaptability to a wide variety of settings and a corresponding diversity of methods and purposes of application. The GAS integrates the development of outcomes scales specifically tailored to the individual, group, agency, or system whose progress is being measured (Kiresuk et al., 1994). When the content of the goals is analyzed it can provide a summary of the intent and success of the individual therapist, patients, and the treatment program. The GAS is able to provide a direct, reliable, and accurate method of assessing the amount of treatment-induced change that has occurred in the client (Kiresuk et al., 1994).

Reliability and validity were researched for the Goal Attainment Scale. Studies that used the 5-point scale reported that inter-rater reliability was high (Kiresuk et al., 1994). The research does suggest that for purposes of validity the GAS score is a 'sensitive measure of change,' (Kiresuk et al., 1994). Ways to improve the reliability and validity of the GAS could include comprehensive training of the raters, adequate definitions of the levels of goal attainment, use of multiple raters, and collaborative goal setting to ensure goals are meaningful and specific to the child's individual needs (Driver, 2006).

GAS Current Research

Rockwood, Joyce, and Stolee (1997) compared outcome measures with GAS to assess patients' sensitivity to changes in health status while undergoing cognitive rehabilitation. The mean gain of the GAS scores were compared to the assessments that were completed prior to the interventions being implemented. Rockwood et al. (1997) reported a large effect size when evaluating responsiveness to change. GAS shows promise as an effective and responsive measure in cognitive rehabilitation (Rockwood et al., 1997).

The GAS is also a responsive method for individual goal setting and treatment evaluation (Steenbeek, Ketelaar, Galama, & Gorter, 2008). In a pediatric rehabilitative setting, the GAS was a useful measure in improving the quality of services that are provided to those children. McDougall and Wright (2009) researched the GAS and integrated the use of the GAS with the International Classification of Functioning, Disability of Health-Child and Youth (ICF-CY). It was established that these tools, used together were able to connect the therapeutic process in a pediatric rehabilitation setting in order to provide consistent clinical care that is family-centered, collaborative, well directed and accountable. The use of the GAS in the pediatric setting facilitates translation of clients' identified needs into distinct, measurable goals set collaboratively by the clients, their families and service providers (McDougall & Wright, 2009). Through the combination of ICF-CY and GAS, it was proven that these were effective measures for patients receiving rehabilitative services. Stokes (2009) found the GAS to be a mathematical technique for quantifying the achievement of goals set, which can be used in rehabilitation. The simplest method for data analysis according to Stokes (2009) is to convert raw scores from the 5-point scale into T-scores which are normally distributed with a mean of 50 and standard deviation of 10.

A study published in the American Journal of Occupational Therapy looked at GAS as a measure of meaningful outcomes for children with sensory integration disorders (Mailloux, Benson, Summers, Miller, Green, Burke, et al., 2007). The study sought to identify recent and

current applications of the GAS as it pertains to children with sensory dysfunction. The first known use of the GAS pertaining to children with sensory dysfunction occurred in two pilot studies that were conducted between 1997 and 2005. This first study used a single-site research for application of the GAS. The second pilot study integrated multiple clinical applications of GAS. Both studies found the GAS to be an effective measure in the medical field with children who have sensory dysfunction. Children were able to make significant gains which were indicated through pretest and posttest results. These studies demonstrated that GAS could capture individual changes in daily life occupations that are functional, meaningful alterations in occupational performance over a short intervention period in a small sample (Mailloux et al., 2007).

GAS and Education

Although the GAS was initially formulated as a means to evaluate individual mental health treatment outcomes, its natural affinity to education manifested itself (Kiresuk et al., 1994). Since the original development of the GAS, there has been additional research which includes a variety of professions including education. A project entitled "Early Identification of Learning Disabilities" (Hegion, Fish & Grace, 1974), consisted of a group of children in grades kindergarten through second who were selected for special education. The children were compared on teacher ratings and standardized tests, and the progress of these children were measured on an individualized basis. There were five specific areas in which goals were developed including reading instructional level, reading level, mathematics instructional level, mathematics level, and vocabulary level. Social and classroom behavior goals were also developed for the children who needed these goals. The students were rated on their progress at the end of the project. A major advantage of the use of the GAS in this project was the facilitation of communication and understanding among the participants (Hegion et al., 1974).

Howe and Fitzgerald (1976) indicated the use of the GAS as part of an evaluation of special education programs in Iowa. There were three levels integrated into the evaluation

process. At the first level, measures which looked at general views held by consumers regarding program and services to children with special needs were applied across the Area Education Agencies (AEA's) and local education agencies. Questionnaires and interviews were also gathered and completed. At the second level, program-specific evaluation was conducted and that information was used to set program goals. Major work priorities, such as Initiation of Pilot Plan, Program Evaluation, and Use of Outcome Data were scaled using the GAS. At the third level, specific behavior and academic skills were targeted and the GAS was used to set individual student goals and monitor improvement in the selected areas. The results of this study indicated that the GAS was an effective measure when conducting a program evaluation as well as student specific goals.

The Student Advocates Inspire Learning (SAIL) project (Balfour & Harris, 1979) used the GAS as part of the development of mandated individualized educational plans in special education programs for emotionally disturbed adolescents. The programs focused on working with and keeping drop-out and drop-out prone students in the mainstream high school experience. Goals were developed using the GAS pertaining to academic and behavior. The goals were reviewed on a weekly basis by the students and staff members to assess the current level of functioning over an 8-week period of time. At the end of that period, the weekly scores were averaged to calculate the final score for the student and show how much progress the student made. The GAS data was an effective method for use with both clinical and administrative purposes: to compare and contrast student subgroups, to provide periodic analyses for the examination of annual trends and for content analyses that allowed analysis of program effectiveness by the type of goal (Kiresuk et al., 1994).

Palmer and Wehmeyer (2002) looked at promoting self-determination in early elementary school students. An approach to teaching self-regulated problem solving and goal-setting skills was used. This was an experimental study being replicated that used the Self-Determined Learning Model of Instruction (Palmer & Wehmeyer, 2002). There were 14 teachers

that participated in this study that had been nominated by administrators. The 50 students in grades kindergarten through third participating in this study were either receiving special education services or were currently in the referral process for special education. The GAS was used and enabled the teachers and students to create goals with predicted and scaled outcomes specific to student need. Interventions were put into place to aid the student in improving. At the end of the study, the teacher and student scales were converted to standardized numeric scores in order to have the data analyzed. The results analyzed included teacher and student paired t-tests, teacher versus grade level chi-square tests, and pre/post-test goals and interest questions paired with sample t-tests. These results indicated that the GAS was effective and showed that more students exceeded expectations than failed to meet or remain the same with their specified goal.

Roach and Elliott (2005) evaluated the GAS as an efficient and effective approach to monitoring student progress. When the GAS is used in educational contexts, teachers, parents, consultants, and students are able to complete the GAS ratings, thus providing an indirect measure of academic or social behavior performance (Roach & Elliott, 2005). In their research, they determined that there has been substantial investigation of the GAS in a variety of mental health and medical settings over the years, but less extensive research and application of the GAS by school psychologists and special educators. There have been other ways that students have had their progress monitored in the educational setting, particularly Curriculum Based Measurements (CBM) and direct observations. By integrating the GAS as a way to develop goals and implement interventions, it provided a more accurate way in gathering and assessing data specific to the student's needs. The GAS has an emphasis on establishing target behaviors and on-going evaluation of academic and behavioral progress which had made this a useful tool for monitoring students' progress and verifying the need for additional support or intervention (Roach & Elliott, 2005). There were case studies reviewed in this article that focused on pre-referral interventions. The results suggest that GAS ratings can provide efficient and accurate assessments of students' academic and behavioral progress. The ease of use for teachers and

other professionals using the GAS is an important advantage. The GAS is perceived as highly acceptable, useful, worthwhile, and valuable for specifying goals and implementing interventions (Roach & Elliott, 2005). It is important that teachers, parents, and professionals using the GAS can understand the development and implementations within it. The GAS promotes clearly operationalized intervention goals and on-going evaluation of student progress, making it a potentially useful tool for special educators and school psychologists (Roach & Elliott, 2005).

Driver (2006) was interested in the GAS as a tool to implement at Darling Point Special Day School where she is the Principal. There was a pilot study conducted with the use of the GAS to determine if it would be an effective measure of the services being provided at the school. After research, she concluded that the GAS is a technique for measuring which goals have been achieved. GAS achieves an individualized, criterion-referenced measure of change, providing a clear expression of plans and outcomes (Driver, 2006). During the study, there were important advantages that were found which include the following: flexibility, relevance, simple and inexpensive, statistically powerful, student and family involvement, collaborative goal setting, acceptability, improved clarity of educational and therapy objectives, improved delivery of interventions and programs, clear line of sight between planning, program delivery, assessment and reporting, more realistic expectations of the educational program, increased levels of satisfaction, and increased motivation towards improvement, provided by the existence of the goals (Driver, 2006). Although there are advantages, some limitations were found by Driver (2006). These include biases in goal scaling and rating, training requirements to implement this approach, and possible temptation to modify goals during the period of the planned intervention (Driver, 2006). The data was gathered and charted to determine which students made progress toward their specified goals, and which students remained the same or worsened. Percentages were calculated and charted using rank of academics to show student achievement. According to Driver, the GAS at Darling Point Special School proved to be an effective measure and provided

teachers, therapists, students and parents with a clear sight from planning to learning and teaching, within a culture of accountability (Driver, 2006).

GAS and Evaluation of School Psychologist Interns

The Ohio Inter-University Council (IUC) of School Psychological Programs incorporates the use of GAS data to evaluate student intern's success at various levels at the three-tier model (Morrison, Barnett & Graden, 2008). Students from Cincinnati University conducted a program evaluation of the Ohio Internship Program in School Psychology to evaluate the effectiveness of interns for the 2007-2008 school year. The focus and 'primary purpose of this state-wide evaluation was to assess the effectiveness and impact of the Ohio Internship Program,' (Morrison et al., 2008). The two components evaluated the demonstration of effectiveness in terms of the intern's competency and skill attainment in the areas that were being evaluated, as well as the demonstration of intern impact on student outcomes. The primary method used for determining outcome was the GAS. The 5-point scale was used with one change to the original GAS. The baseline goal which is named "Expected Level of Outcome" was replaced with "No Change" in order to 'better represent students' responses to the intervention,' (Morrison et al., 2008). The other rating scales remained the same and include: somewhat more than expected, somewhat less than expected, much more than expected, and much less than expected.

Interns were asked to provide GAS outcome data for six individual, targeted, and universal interventions which included three academic interventions and three behavior interventions. Once data were collected with the GAS, summary statistics to quantify intervention outcomes were calculated (Morrison, et al., 2008). These summary statistics included the percentage of non-overlapping data (PND) and effect size (ES). GAS outcomes were gathered for 567 interventions provided by school psychology interns to Ohio students. The interventions included 282 academic interventions and 285 behavior interventions. Data was analyzed using PND and ES. The results suggest strong impact of the intervention services that were provided to the students in Ohio by the school psychology interns. The PND was found to

be mildly to moderately effective, and the ES resulted in a large effect which indicates that these students made significant gains.

The GAS is now utilized in the Marshall University Summer Enrichment Program as a way of determining effectiveness of the services and interventions provided by the summer practicum students.

Marshall University Summer Enrichment Program

The Marshall University Summer Enrichment Program (MUSEP) is a hands-on lab school which provides an opportunity for graduate students to apply and demonstrate the knowledge and skills they have gained during coursework. There are different disciplines participating in the program which consist of School Psychologists, School Counselors, Literacy Specialists, and Special Education Teachers. Graduate students from each discipline are assigned to a multidisciplinary team by a program director. The team assignment determines what ages of students the members will be working with. There is a three hour orientation where the teams are introduced and are provided with an overview of the summer program, goals, and objectives, as well as engage in team-building activities.

An essential part of the summer program is team collaboration and this is a focus during the first week prior to the children arriving. The graduate students are trained and work with their formed team on teambuilding, collaboration, and diagnostic teaching of reading through short cycle assessment and curriculum based-assessment. All students participate in evaluating the effectiveness of the program, the impact of the program on school children, and the satisfaction of parent consumers (Krieg, Miekamp, O'Keefe, & Stroebel, 2006). Each team is responsible for developing a portfolio of their work including lesson plans, assessment data, evaluation of the student's progress, and program success. Therefore, it is imperative that these teams work collaboratively to reach their goals (Krieg, et al., 2006).

The children arrive on the second week for instruction from Monday through Thursday. The instructional day is from 7:30am to 12:30pm. Literacy is at the center of the curriculum,

evidenced by an uninterrupted reading block each day (Krieg, et al., 2006). All team members must participate in the instruction of the 90 minute reading block which focuses on the needs of the child. Assessment information on each child is established and from there, put into groups based on his or her instructional needs and skill level.

Observations are provided by site supervisors who provide critique and feedback on graduate student performance and competency in the areas being assessed. The supervisors meet with the graduate students on a daily basis to discuss any concerns that may arise and help to implement problem solving strategies. At the conclusion of each day, the graduate students attend a group supervision session to discuss the day's events.

Marshall University Evaluation of School Psychologists

The School Psychology graduate students are evaluated during the summer school program. The 16 areas that are evaluated are pertinent to the field of school psychology and give the graduate students opportunities to apply and demonstrate the skills they have learned in their coursework and practicum placements. The school psychology graduate students are observed and rated on the skills they are demonstrating at the time and rated on a scale from 1 – 5 or Not Observed (NO). The scale includes the following: 1 - Unsatisfactory; 2 – Needs Improvement; 3 – Satisfactory; 4 – Very Good; and 5 – Excellent. The graduate students must earn a 3 or higher in order to obtain competency in that area. If a student receives an NO, 1, or 2 then they must be able to demonstrate that skill when they are observed again.

At the end of the six week summer program, an overall evaluation is completed with the student. Scores from the observations are used to evaluate the students in the following areas: Participation, Data Based Decision Making, Counseling, Developmental Guidance, Behavior Modification, Consultation and Teaming, Parent Communication, and Reception to Supervision. The graduate student receives either a rating of Superior Evidence, Adequate Evidence, or Insufficient Evidence.

Students are also evaluated by examining outcome data collected with the GAS. This tool has proven to be effective and aids interns and practicum students in determining if the services and interventions they are providing are helping students improve. Throughout the year, interns and practicum students are able to assess their effectiveness in the services they provide by monitoring student outcome data.

The current study utilizes the GAS data to determine the effectiveness of the school psychology practicum students and the MUSEP.

Statement of Hypotheses

1. On average, children will improve.
2. On average, children will improve academically.
3. On average, children will improve behaviorally.

CHAPTER II

Method

Participants

The participants in this study were 11 Marshall University School Psychology graduate students. They worked in teams of graduate students from Special Education, School Counseling, and Literacy. Students were divided into seven teams during the 2008 Marshall University Summer Enrichment Program. Team members rated 128 children on academic and/or behavior goals.

Procedure

The School Psychology Practicum Students at Marshall University Summer Enrichment Program evaluate their effectiveness by using the Goal Attainment Scale. The GAS has a -2 to +2 rating depending on the progress of the goal that is developed. The children start at 0 which is the expected level of outcome. If a child does not improve, the rating can be -1 which indicates a somewhat less expected level of outcome, or -2 which indicates a much less expected level of outcome. If the child stays the same, the rating can be 0 which indicates no change. If the child does make progress, the rating can be +1 which indicates a somewhat more expected level of outcome, or +2 which indicates a much more expected level of outcome.

During the 2008 Summer Enrichment Program, each team developed a goal for each child specific to his or her need. All children in the summer school program had at least one goal. The 11 School Psychology students were the lead evaluators and educated their teammates which consisted of school counselors, literacy specialists, and special education teachers, on how to develop goals for each child. All goals were geared specifically to each child's individual need and consisted of academic, behavior, or both.

The Ohio School Psychology Internship Program has developed a Step-by-Step Guide to Developing and Scaling Goals Using the Goal Attainment Scaling. This is the same scaling

system that the Marshall University school psychologist practicum students utilized to develop specific goals for children. The steps are as follows:

STEP 1 – Specify the Expected Level of Outcome for the Goal

> As part of the problem-solving process, you will develop a goal statement that is observable, measurable, and specific. Goals should be based on baseline data, goals should be realistically ambitious, based upon what the student will likely achieve by the end of the intervention, goals should take into consideration the usual outcomes of this intervention, the resources of the student, the amount of time planned for the intervention, and the skills of the intervention specialist/change agent, goals should be socially valid (i.e., acceptable to teachers, parents, and the student) and goals should be stated in the positive (i.e., promoting replacement behaviors).

Step 2 – Review the Expected Level of Outcome given the following considerations

> *Relevance*: Is the goal relevant to the student's present situation?
> *Availability*: Are the intervention services necessary to attain this goal available?
> *Scale Realism*: Is the expected level of outcome realistic for this student at this time with this intervention?

Step 3 – Specify the Somewhat More and Somewhat Less Than Expected Levels of Outcome for the Goal

> Provide observable, measurable descriptions of outcomes that are more or less favorable than the expected outcomes in the boxes immediately below and immediately above, respectively. These descriptions are less likely to occur for this student, but still represent reasonably attainable outcomes.

Step 4 – Specify the Much More and Much Less Than Expected Levels of Outcome

> Complete the extreme levels of the scale with descriptions of the indicators that are "much more" and "much less" favorable outcomes that can be realistically envisioned for the student. Each extreme level represents the outcome that might be expected to occur in 5% to 10% of similar at-risk students (Morrison, et al., 2008).

Baseline data for each child were collected and a goal was developed. Interventions were implemented over a five week period then performance data were collected. Children were rated on a scale of -2 to +2 depending on the progress they made towards the specific goal. The Summer Enrichment Program had a total of 128 children enrolled who had goals that were developed, implemented, and rated over a 5 week period of time using the Goal Attainment Scale.

CHAPTER III

Results

Of the 128 children, 89 children had academic goals only, 27 children had behavior goals only, and 12 children had both academic and behavior goals. The collected data were analyzed for academic goals and behavior goals. A paired t-test was utilized in this study to analyze nonparametric data. This method had been established as an effective strategy for distributions with small numbers of values (Snedecor & Cochran, 1967). A t-test was run on the scores of all children to determine if the graduate students were effective in helping children to improve. Results indicated $t = 8.364$ with a mean difference of .86719 and a standard error mean of .10368 ($p<.001$). This suggests that on average, children improved and the graduate students were effective in providing services and interventions. The t-test was then calculated for academic and behavior goals separately. For academics, $t = 7.147$ with a mean difference of .8910 and a standard error mean of .12015 ($p<.001$). For behavior, $t = 4.903$ with a mean difference of .7750 and a standard error mean of .15806 ($p<.001$). These results indicated that, on average, the children improved on academic and behavior goals.

A Univariate Analysis of Variance (UNIANOVA) was run to determine variation within groups, as well as the variation between groups in looking at both academic and behavior goals. The two dependent variables were academics and the behaviors which were both rated on 5-point scales. Teams One through Seven was the independent variable. When the data were analyzed for differences among the seven teams, there were no significant differences because $p = .140$ for academics and $p = .109$ for behavior. Since scores were not normally distributed, requirements were not met for using parametric statistics (Levin & Fox, 2007). During the analysis, not all children had both an academic and behavior goal which made the size of N extremely small for behavior goals. This made it difficult to obtain a normal distribution with the data that was available. Because of these reasons it was determined that the data should be analyzed using a nonparametric statistic which 'does not require normality or the interval level of measurement,'

(Levin & Fox, 2007). The data were analyzed using a ranking statistical method known as the Kruskal-Wallis Test. In this case, the dependent variables were the rank of academic scores and rank of behavior scores. Table 1 shows the results of this analysis.

Table 1
Kruskal-Wallis Test

	Academics	Behavior
Chi-Square	12.274	8.725
df	6	6
Significance	.056	.190

Table 2 shows the mean ranks of academics for the Kruskal-Wallis test.

Table 2
Rank of Academics

Team Number	N	Mean Rank
1	11	68.00
2	12	45.54
3	12	39.38
4	16	38.66
5	20	59.93
6	13	48.92
7	17	54.76
Total	101	

Table 3 shows the mean ranks of behaviors for the Kruskal-Wallis test.

Table 3
Rank of Behaviors

Team Number	N	Mean Rank
1	5	15.60
2	2	29.25
3	4	21.00
4	4	18.50
5	10	20.70
6	9	26.83
7	6	12.83
Total	40	

The results of the rankings indicate that there is no significant difference between the groups for academic or behavior goals suggesting that the practicum students were equally effective in the services and interventions that were provided.

Frequencies of academics on the 5-point rating scale are shown in Table 4.

Table 4
Academics

Score	Frequency	Valid Percent
-2	6	5.9
-1	7	6.9
0	22	34.7
+1	23	57.4
+2	43	42.6
Total	101	100.0

The teams were effective in helping children improve their academic goals when looking at ratings +1 and +2. There were 23 children that improved at a somewhat more than expected level. The graduate students were able to help 57.4% of those children improve their academic goal. There were 43 children that improved at a much more than expected level. The graduate students were able to help 42.6% of those children improve to the best of their ability.

Frequencies of behaviors on the 5-point rating scale are shown in Table 5.

Table 5
Behavior

Score	Frequency	Valid Percent
-2	2	5.0
-1	2	5.0
0	7	17.5
+1	21	52.5
+2	8	20.0
Total	40	100.0

The teams were effective in helping the children improve their behavior goals when looking at ratings +1 and +2. There were 21 children with behavior goals that were rated at +1. Graduate students were able to help 52.5% of the children improve their behavior goal at a somewhat more expected level of outcome. There were 8 children with behavior goals that were rated at +2. Graduate students were able to help 20% of those children reach a much more than expected level of outcome. Thus, effectiveness of graduate students was demonstrated as these children made gains in their behavior.

To further determine whether the services and interventions provided by the graduate students were effective, the percentage of nonoverlapping data points (PND) was calculated. The PND was assessed by adding the total number of +1 and +2 scores for academic and behavior goals, and then dividing that number by the total number of academic and behavior goals. The results of the academic goals revealed a PND of 65% which is mildly effective when compared to the total number of academic goals. The results of the behavior goals revealed a PND of 72% which is moderately effective when compared to the total number of behavior goals. When scores of academic goals and behavior goals were calculated together, results revealed a PND of 74% which is moderately effective.

CHAPTER IV

Discussion

The purpose of this study was to evaluate the effectiveness of school psychology graduate students in the MUSEP using the GAS to measure outcome data. The GAS was used to evaluate academic and behavior goals of the children in the summer program.

It was hypothesized that children would improve after participation in the MUSEP. A t-test was conducted and results suggested that on average, children improved on academic and behavior goals. A Kruskal-Wallis Test revealed that the graduate students on different teams were equally effective in the interventions and services they provided. Children at all age levels benefited from the interventions provided.

The GAS has proven to be an effective tool in measuring student outcome data as well as evaluating the services provided by professionals. This study, as has been shown in previous studies, (Balfour & Harris, 1979; Driver, 2006; Hegion et al., 1974; Howe & Fitzgerald, 1976; Palmer & Wehmeyer, 2002; and Roach & Elliott, 2005) demonstrates that GAS is a useful tool in the educational setting. The current study and the Ohio IUC study (Morrison et al., 2008) both integrate the use of the GAS as an outcome measure for the services and interventions provided by school psychologists. Although, IUC includes the effectiveness of interns throughout the year and over a broader variety of topics that are not limited to academic and behavior interventions. For Ohio interns, the GAS is used to determine the effectiveness of consultation services, crisis management, systems level interventions, prevention, and academic and behavior interventions (Morrison et al., 2008). Another difference is that MUSEP is evaluating school psychologists' effectiveness in working with teams whereas Ohio IUC is assessing the effectiveness of school psychologists working independently. Even with these differences, both studies revealed that graduate students were mildly to moderately effective in the services and interventions that were provided when the PND was calculated. These studies revealed the consistency of application of the GAS as a measure of outcome.

This study is comparable to Palmer and Wehmeyer (2002) in the way that the data were analyzed. Both studies used a paired t-test to configure results which revealed that on average, children did improve in their goals. In both studies, more children were able to exceed expectations. The GAS was a consistent and effective outcome measure. In looking at the current study and Roach and Elliott (2005), the GAS proved to be an effective measure and provided an accurate way to gather and assess data specific to the children's needs. Both studies incorporate behavior and academic goals and monitor progress on each. Driver (2006) incorporated the use of the GAS with children's academic achievement. The MUSEP also used the GAS as a way to measure children's academic achievement. Both studies calculated the percentage of children who improved on academic goals, and used rank of academic to show children's achievement.

The GAS has proven to be an effective tool across various disciplines. It has been used as a way to measure outcome in a variety of settings, as well as across a variety of ages. The results of this study suggest that the practicum students were effective in helping children improve academically and behaviorally.

Limitations

A limitation to this study is the absence of a control group which eliminates a comparison of children who did not receive services. Although there is no control group, when comparing with Ohio interns (Morrison et al., 2008), this study duplicates a study with a similar design and found the GAS an effective tool. Another limitation is the fact that there was a small sample size of behavior goals for children which limited the data analysis to a nonparametric measure. A final limitation is the limited geographical region and population. This study may not apply to children from other regions.

Recommendations

It is recommended that further study be conducted with the GAS as a measure to evaluate effectiveness pertaining to the MUSEP. This is the first study evaluating the GAS at MUSEP and for future studies, all children in the MUSEP could have both an academic and a behavior goal. This would enable the study to have more data which would create a bigger sample population. In turn, the actual collected data could be analyzed and compared to the GAS scaling to give more insight into the effectiveness of the GAS for showing gains. This would enable analysis in more ways than just the nonparametric method which is what this study used.

A random sample may enable a more accurate outcome of assessment when monitoring progress on academic goals. For example, if a child has a reading goal of 50 words per minute at the expected level of outcome and needs to progress to between 60 and 70 words per minute for improvement, that child could be called on randomly to read. Other children with reading goals would randomly be called on as well. Data would be collected through the random sample and progress would then be monitored and rated.

Data could also be collected by using the actual goals developed for each child. The goals could be grouped into categories of the specified goal, for example, reading goals, writing goals, behavior goals, etc. This would enable more data, which could provide the sample with a larger N. More methods of data analysis could then be used to determine if services and intervention were effective.

References

Allison, R. (2002). Best Practices in Supervision of School Psychology Staff. In A. Thomas, & J. Grimes (Eds.), *Best Practices in School Psychology IV* (pp.122-123). Bethesda, MD: National Association of School Psychologists.

Balfour, M.J., & Harris, L.H. (1979). *SAIL, an innovative mainstream special education project for drop-out prone students: Validation report.* Hopkins, MN.

Chafouleas, S.M., Clonan, S.M., & Vanauken, T.L. (2002). A National Survey of Current Supervision and Evaluation Practices of School Psychologists. *Psychology in the Schools, 39(3),* 317-324.

Driver, C. (2006). *Goal Attainment Scaling, Inclusive Education.* Retrieved July 20, 2009 from the Learning Place website: http://www.learningplace.com

Fuchs, L. (2002). Best Practices in Defining Student Goals and Outcomes. In A. Thomas, & J. Grimes (Eds.), *Best Practices in School Psychology IV* (pp.558). Bethesda, MD: National Association of School Psychologists.

Gibbons, K., & Shinn, M. (2002). Best Practices in Evaluating Psychoeducational Services Based on Student Outcome Data. In A. Thomas, & J. Grimes (Eds.), *Best Practices in School Psychology IV* (pp.159-161, 166). Bethesda, MD: National Association of School Psychologists.

Gibbons, K., & Silberglitt, B. (2008). Best Practices in Evaluating Psychoeducational Services Based on Student Outcome Data. In A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology V* (pp.2104). Bethesda, MD: National Association of School Psychologists.

Hegion, A., Fish, E., & Grace, S. (1974). Goal Attainment Scaling: An alternative technique for measuring academic progress. *Goal Attainment Review, 1,* 20-34.

Howe, C.E., & Fitzgerald, M.E. (1976). *A model for evaluating special education programs in Iowa area education agencies.* Des Moines, IA: Iowa State Department of Public Instruction, Division of Special Education.

Kiresuk, T.J., Smith, A., & Cardillo, J.E. (1994). *Goal Attainment Scaling: Applications, Theory, And Measurement.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Kiresuk, T.J., & Sherman, R.E. (1968). Goal attainment scaling: A general method for evaluating comprehensive community mental health programs. Community Mental Health Journal. 4(6), 443-453.

Krieg, F.J., Meikamp, J., O'Keefe, S., & Stroebel, S.S. (2006). Field-based Experience in Light of Changing Demographics. *Trainer' Forum*, 25, 4-7.

Levin, J., & Fox, J. (2007). *Elementary Statistics in Social Research.* Boston, MA: Pearson Education, Inc.

Mailloux, Z., Benson, T.A., Summers, C.A., Miller, L.J., Green, B., Burke, J.P., et al. (2007). Goal Attainment Scaling as a measure of meaningful outcomes for children with sensory integration disorders. *The American Journal of Occupational Therapy, 61(2),* 254-258.

McDougall, J., & Wright, V. (2009). The ICF-CY and GAS: Benefits of their combined use for pediatric practice. *Disability and Rehabilitation, 31(16),* 1362-1372.

Morrison, J.Q., Barnett, D.W., & Graden, J.L. (2008). Evaluation of the Ohio Internship Program in School Psychology. University of Cincinnati.

National Association of School Psychologists (2000). *Position Statement on Supervision in School Psychology.* Retrieved July 12, 2009, from the NASP website: http://www.nasponline.org/about_nasp/pp_supervision.aspx

National Association of School Psychologists (2004). *Standards for training and field placement programs in school psychology.* Retrieved July 12, 2009 from the NASP website: http://www.nasponline.org.

Palmer, S.B., & Wehmeyer, M.L. (2002). Promoting self-determination in early elementary School: Teaching self-regulated problem-solving and goal setting skills. *Remedial and Special Education, 24,* 115-118.

Roach, A.T., & Elliott, S.N. (2005). Goal Attainment Scaling: An Efficient and Effective Approach to Monitoring Student Progress. *Teaching Exceptional Children, 37(4),* 8-16.

Rockwood, K., Joyce, B., & Stolee, P. (1997). Use of goal attainment scaling in measuring clinically important changes in cognitive rehabilitation patients. *Journal of Clinical Epidemiology, 50(5),* 581-588.

Snedecor, G.W., & Cochran, W.G. (1967). *Statistical methods.* (6th ed.) Ames, IA: The Iowa State University Press.

Steenbeek, D., Ketelaar, M., Galama, K., & Gorter, J.W. (2008). Goal Attainment Scaling in Pediatric Rehabilitation. *Child: Care, Health and Development, 34(4),* 521-529.

Stokes, L. (2009). Goal Attainment Scaling in Rehabilitation. *Clinical Rehabilitation, 23(4),* 362-370.